

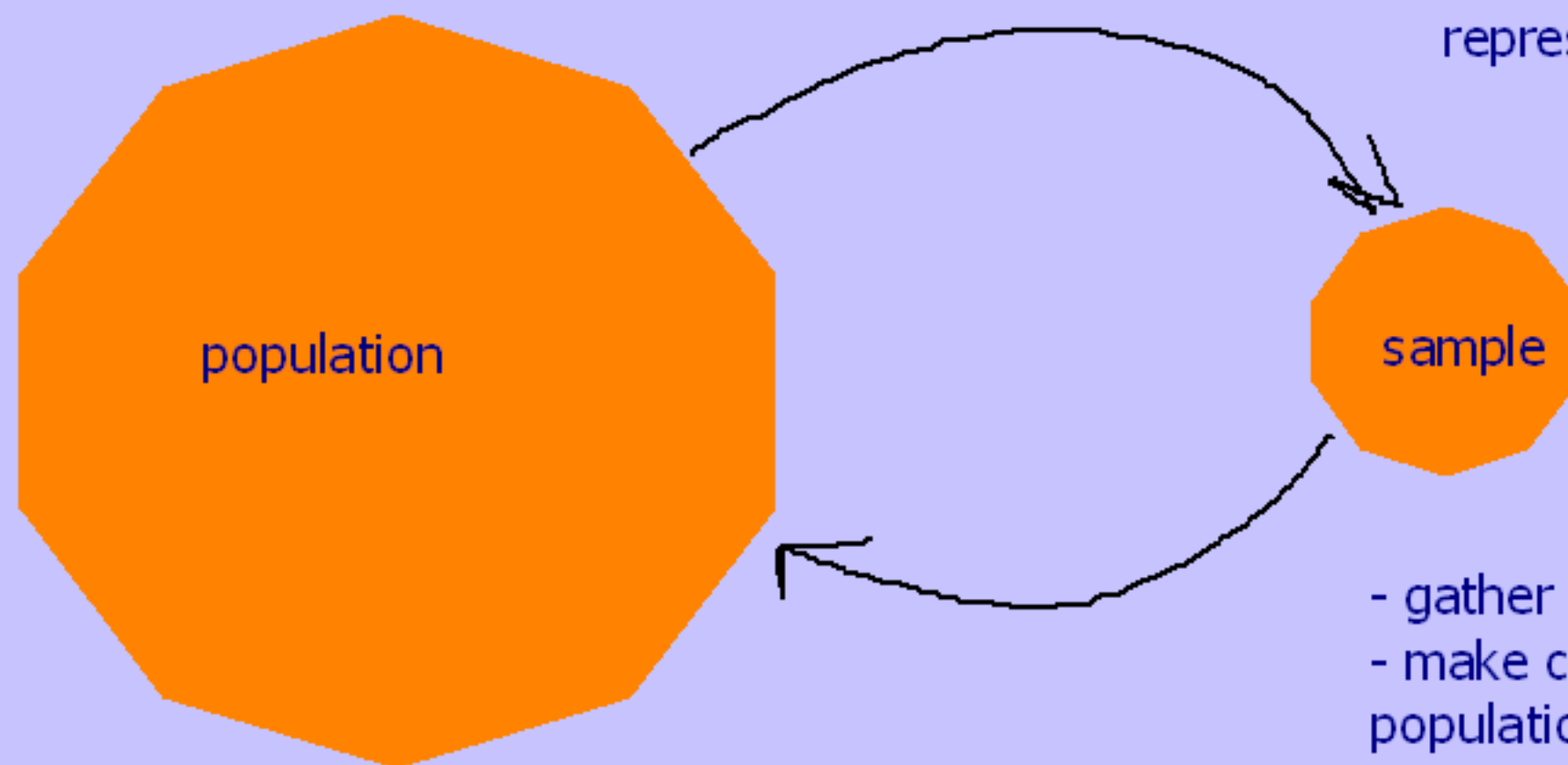
Some Introductory Statistics Vocabulary...

Population versus Sample

Population - All of a certain group

Sample

- A small section of the population
- can be good or bad representation of the pop.



- gather data from sample
- make conclusions about population

Data.....what does it consist of?

Individuals:

The objects described by a set of data

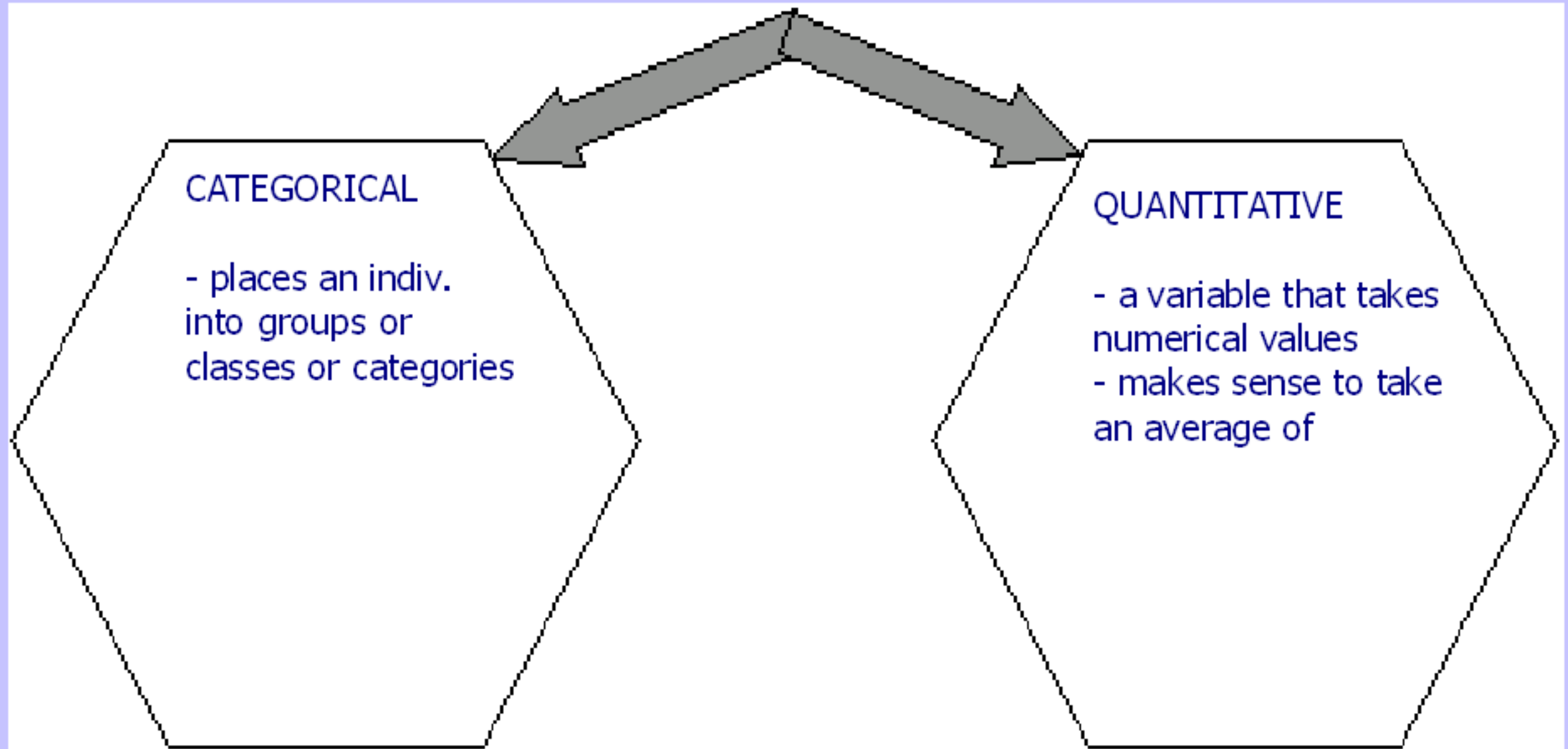
Ex: Colleges, cars, students, plants, dogs, etc.



Variables:

- characteristic of an individual
- something we measure about the individual
- different values for different individuals

Variables:

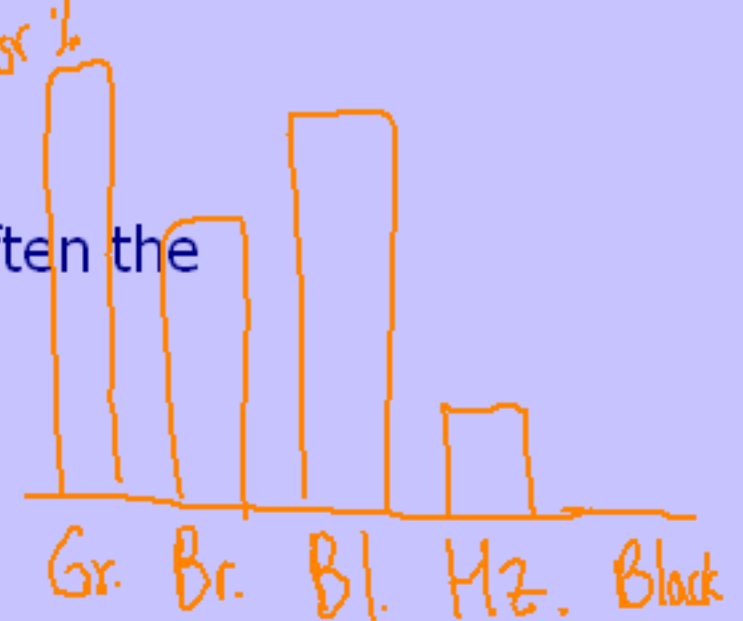


EXAMPLE: Variables about college

class size	avg.	Q
majors	#	Q
location		C
distance,	miles from home	Q
tuition	\$\$\$\$	Q
clubs/sports	#	Q
private/public		C
division		C
faculty mem.		Q

Distribution (of a variable)-

- Shows all the different values of the variable and how often the variable takes those values
- Examples: histogram, stemplot, bar chart, etc.



Exploratory Data Analysis-

- Uses graphs, distributions, numerical summaries to describe a set of data
- We will do this for Ch. 1-5 in the book

averages, Med.
min, max
std. dev %

* **SECOND PART OF THE COURSE: INFERENCE
Inference (to infer something)-

Ch. 6-10

educated guess based on data w/ certainty

Input your data on the side board by putting a tally mark next to where you fit

HAIR COLOR

Blond

Brown

Black

Red

Mixed

EYE COLOR

Blue

Green

Brown

Hazel

Mixed

Other

ZIP CODE

18976

18901

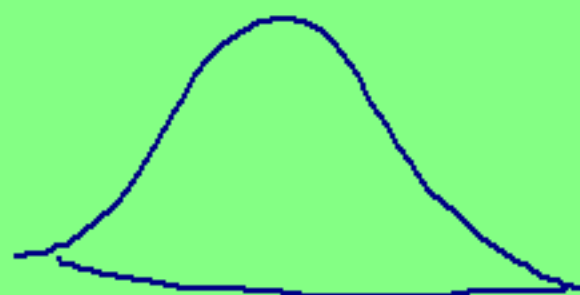
Other

AP Stat
Section 1.1- Notes

Graphs/Distributions:

Which ones will we do?

- **Bar chart**
- **Pie chart**
- **stem plot**
- **dotplot**
- **histogram**
- **boxplot**
- **smooth curves**



Categorical Distributions:

1. Bar Graph

X-Axis => categories/values of the variable

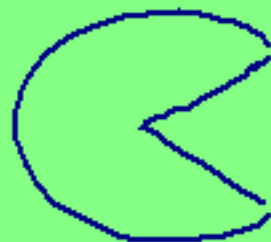
Y-Axis => count = # = frequency

OR

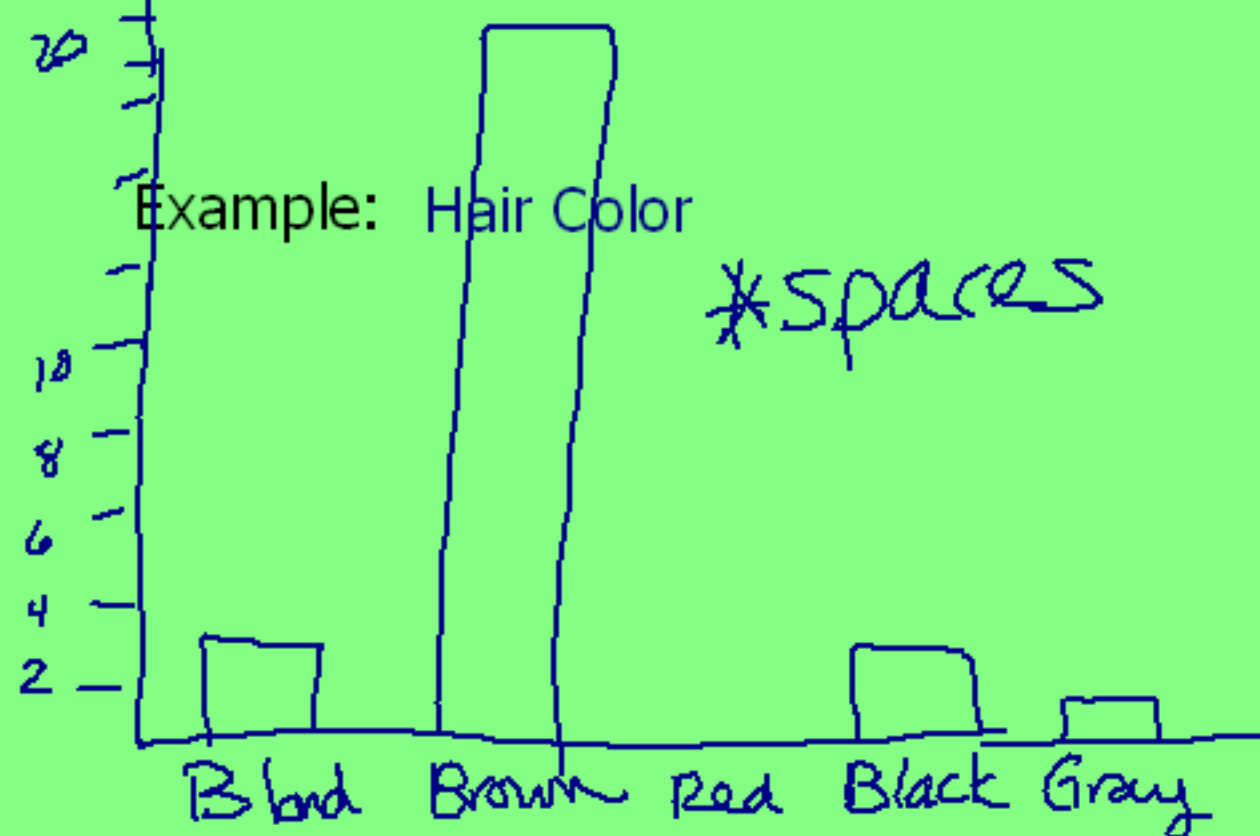
% = relative frequency

freq.
#

2. Pie Chart



- shows all parts of a whole
- shows only % or relative frequency
- find % of each category, then sketch graph



Example: Hair Color

Eye Color



Quantitative Distributions:

1. Stemplot (aka Stem and Leaf Plot)

- Separate... each observation into a stem and leaf last digit
- Write... stems in order (small to large) vertically, with a line to the right
- Write... leaves in the row to the right of the stem in order

Example 1:

Babe Ruth's homerun totals each season for the Yankees:

~~54, 59, 35, 41, 46, 23, 47, 60, 54, 46, 49, 41, 34, 22~~ 102

Create a stemplot:

Babe Ruth
HRs

2		2	5
3		4	5
4		1	1 6 6 7 9
5		4	4 9
6		0	

102
115
009

Example 1: Babe Ruth's homerun totals each season for the Yankees:

~~54, 59, 35, 41, 46, 25, 47, 60, 54, 46, 49, 41, 34, 22~~

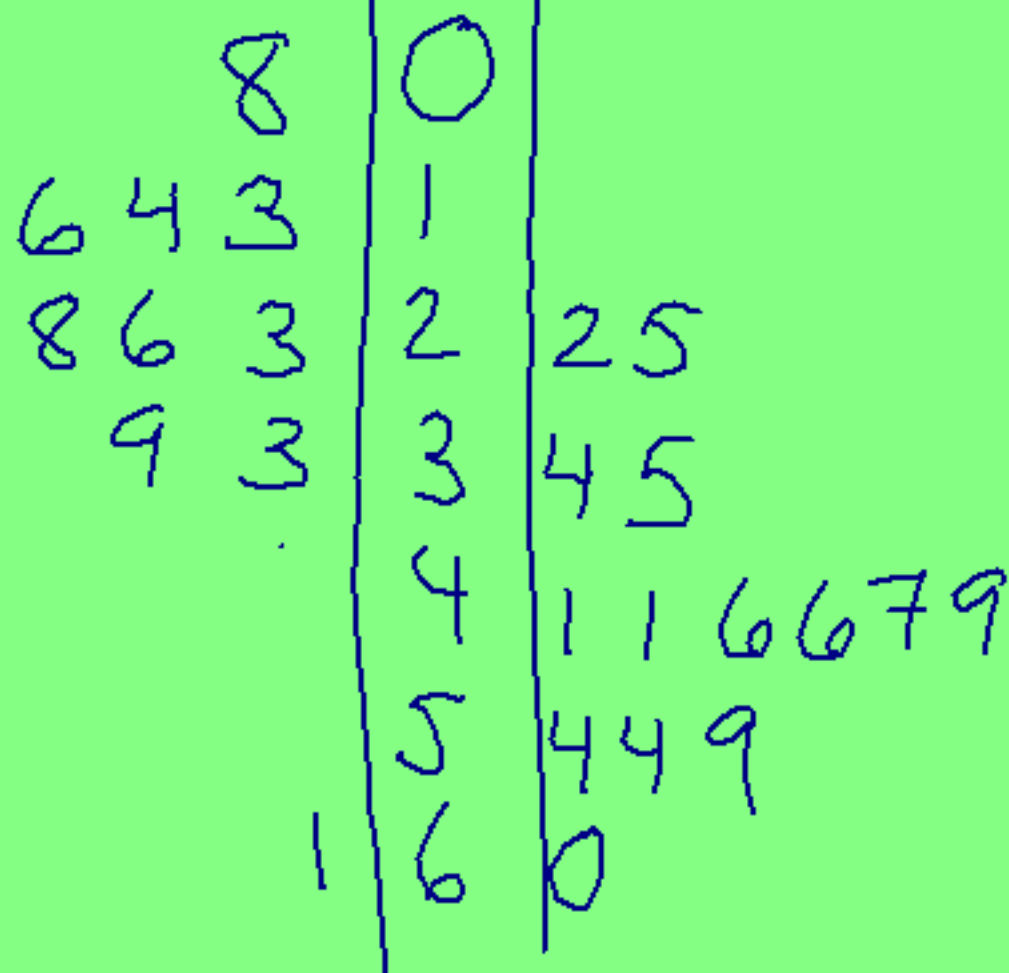
Example 2: Roger Maris' homerun totals for the Yankees:

~~8, 13, 23, 33, 28, 16, 14, 39, 26, 61~~

Create a Back-to-Back stemplot

Roger Maris

Babe Ruth



Splitting Stems:

When? when observations are clumped on only one or a few stems

In what ways can the stems be split?

- 2 stems for every 10 digits
- 5 stems for every 10 digits
- 10 stems for every 10 digits

Example 3: Age guesses 1 stem for every 10 digits:

BLOCK 1 & 2

[illegible]

2 stems for every 10 digits:

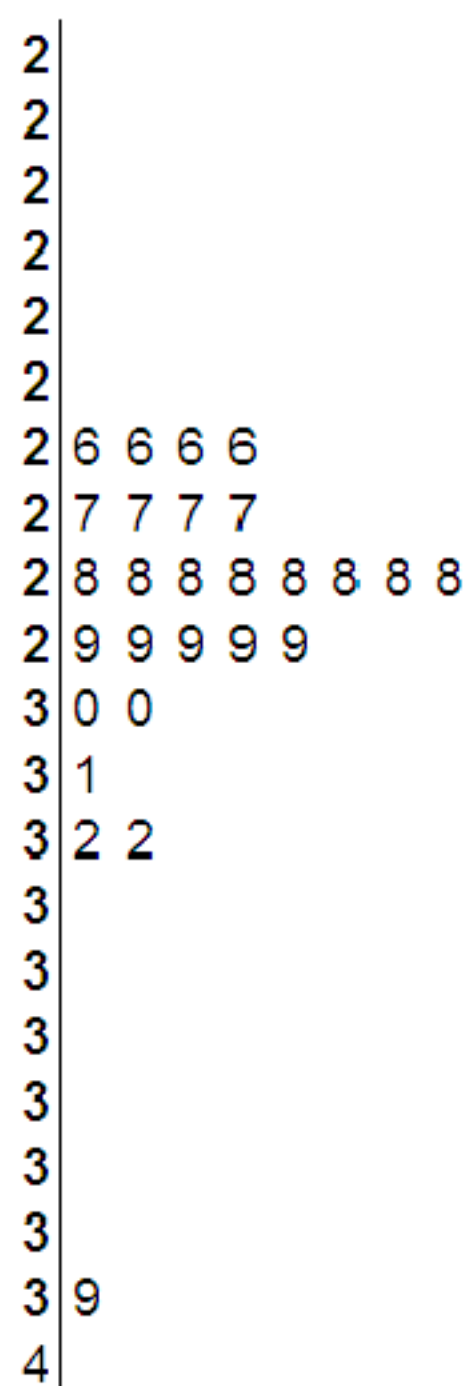
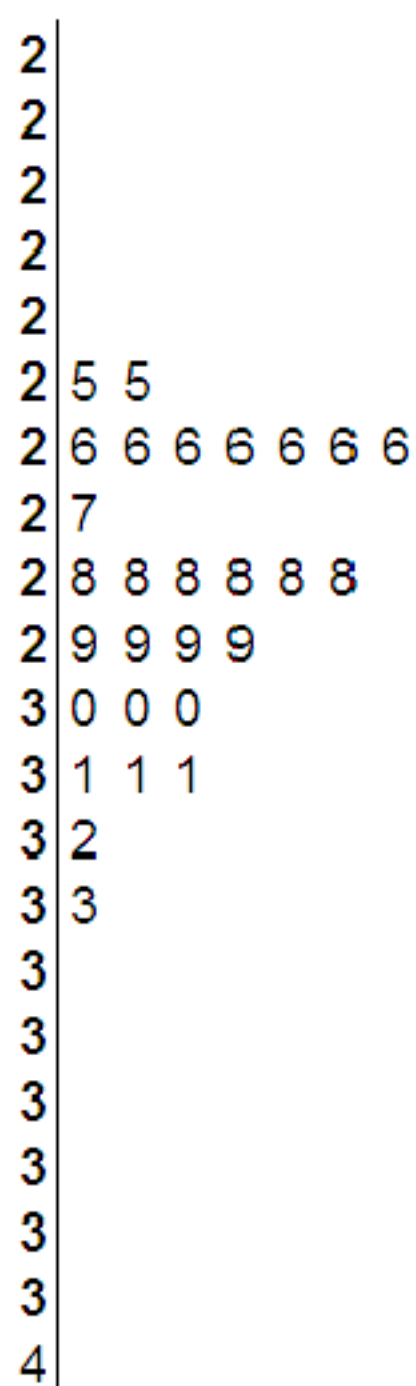
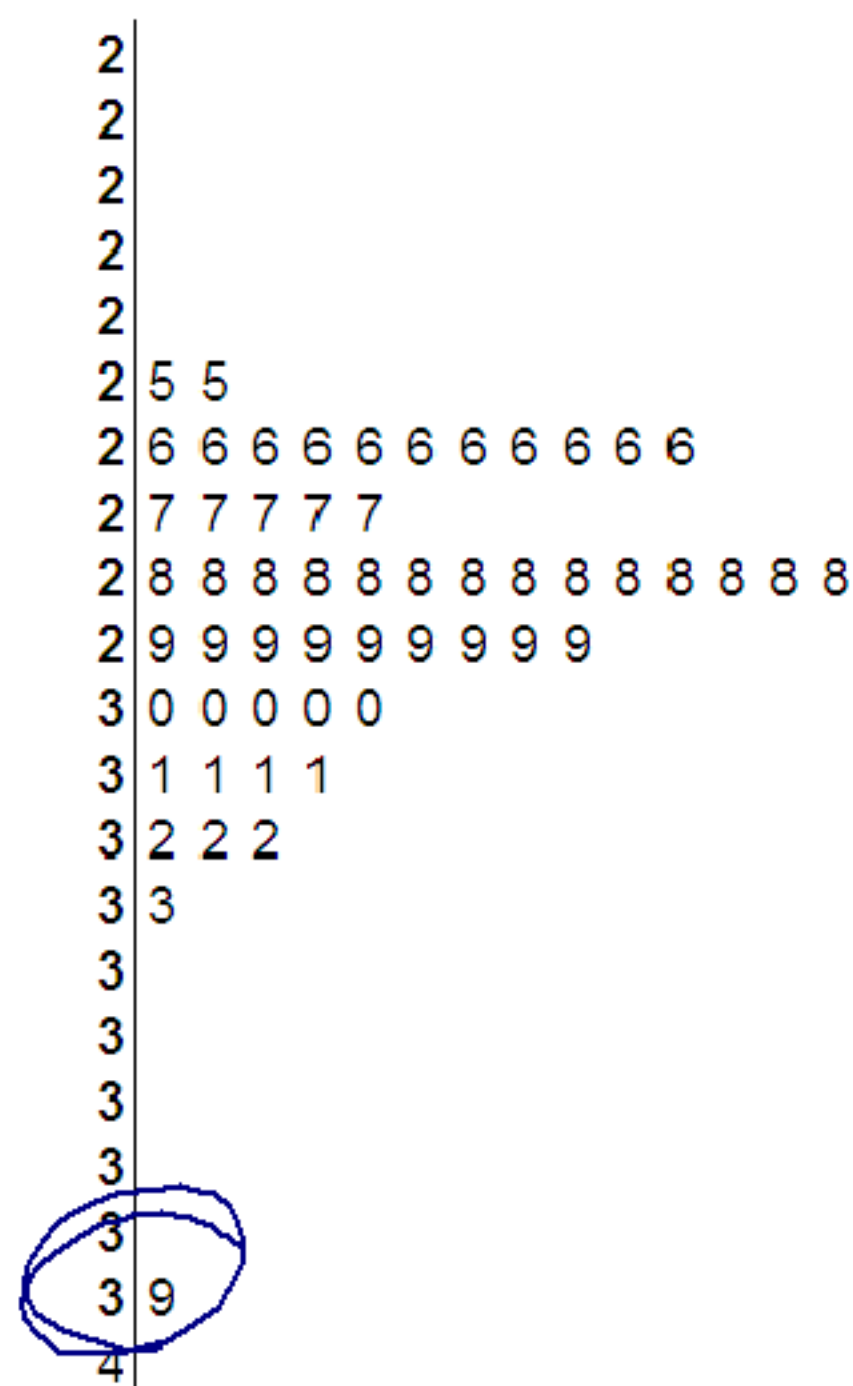
[illegible]

Age guesses 1 stem for every 10 digits:

Block 1 & 2

Block 1

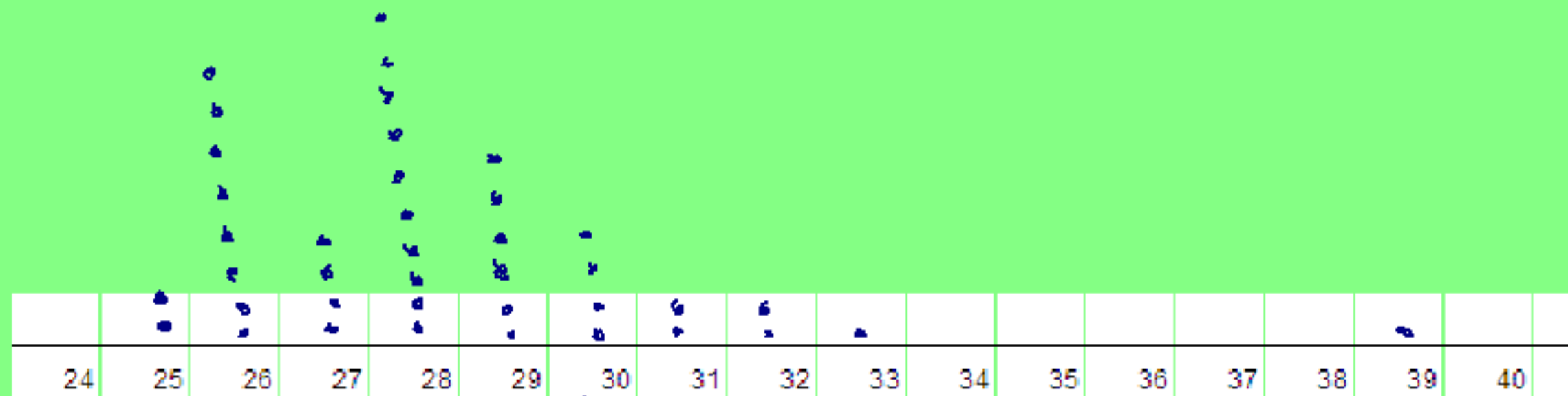
Block 2



2. Dotplot

- Each... value of the variable is placed on the horizontal axis
- *Every time...* you observe a value, you put a dot above it on the plot in a vertical line
- *Better for...* smaller sets of data

Example 1: Age Guesses

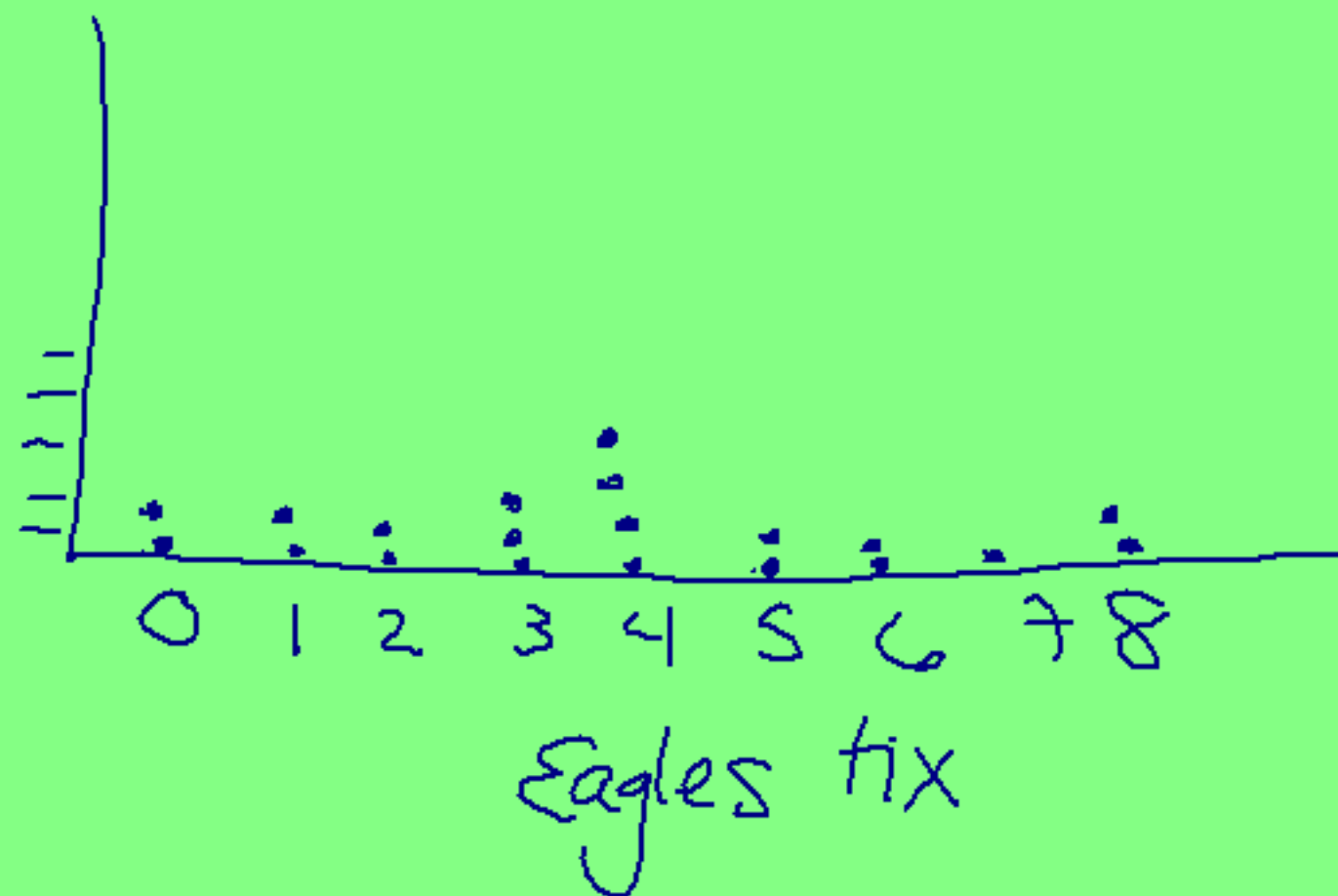


Age Guesses

Example 4: Below are the number of Eagles tickets bought per buyer during a 1 hour period on ticketmaster.com:

~~8, 5, 0, 7, 1, 4, 4, 0, 2, 1, 4, 5, 3, 0, 8, 4, 3, 2, 6, 3~~

Create a dotplot



- Please take a look at the sheet entitled:
Inputting and working with lists
- Please find a partner, get out your calculators, and
come get a link off the front table
- Please complete the worksheet

3. Histogram

- *Take...* data and split into equal classes/groups
 - o *Understand* your data
 - o *Choose* logical classes
 - o *Classes* can't overlap
- *Count...* # of observations in each class. This is the height of the bar for that class
- *Create...* bars for each class
 - o *X-axis:* classes/groups
 - o *Y-axis:*

- frequency	- cumulative frequency
- relative freq.	- cumulative relative freq.

Example: page 14, example problem 1.9

Example 2: Test Scores (LTEST)

 Create a frequency histogram on the calculator

Creating other types of histograms:

Using a chart is very helpful when creating the other types of histograms:

* Use the list TEST that you transferred

Class/Group	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Freq.
Total:				

Graphs:

**** Complete the histogram example worksheet**

Comparison:

Stemplots

- intervals determined by our # system
- shows all data points
- tedious for large data sets
- shows only frequency

Histograms

- classes determined logically
- doesn't show all data points
- easier to create
- shows both freq. and relative freq.

